

SOLVABILITY OF SECOND-ORDER EQUATIONS WITH HIERARCHICALLY PARTIALLY BMO COEFFICIENTS

HONGJIE DONG

ABSTRACT. By using some recent results for divergence form equations obtained in [7, 10], we study the L_p -solvability of second-order elliptic and parabolic equations in nondivergence form for any $p \in (1, \infty)$. The leading coefficients are assumed to be in locally BMO spaces with suitably small BMO seminorms. We not only extend several previous results by Krylov and Kim [14]-[18] to the full range of p , but also deal with equations with more general coefficients.

1. INTRODUCTION

In this paper, we study the $W_p^{1,2}$ -solvability of parabolic equations in nondivergence form:

$$Pu - \lambda u = f, \quad (1.1)$$

where $\lambda \geq 0$ is a constant, $f \in L_p$, and

$$Pu = -u_t + a^{jk} D_{jk} u + b^j D_j u + cu.$$

We assume that all the coefficients are bounded and measurable, and a^{jk} are uniformly elliptic, i.e., for some $K > 0$ and $\delta \in (0, 1]$,

$$|b^j| + |c| \leq K, \quad |a^{jk}| \leq \delta, \quad \delta |\xi|^2 \leq a^{jk} \xi^j \xi^k \leq \delta^{-1} |\xi|^2.$$

If all the coefficients are time-independent, we also consider the W_p^2 -solvability of elliptic equations in nondivergence form:

$$Lu - \lambda u = f, \quad (1.2)$$

where

$$Lu = a^{jk} D_{jk} u + b^j D_j u + cu.$$

The main purpose of this paper is to show how the recent results obtained in [7, 10] for *divergence form* equations can be used to i) overcome the restriction $p > 2$ in the L_p -theory of *nondivergence form* equations with partially BMO coefficients developed by Krylov and Kim in [14]-[18]; ii) extend the results to systems as well as equations with a more general class of coefficients.

2010 *Mathematics Subject Classification.* 35J15, 35K15, 35R05.

Key words and phrases. Second-order equations, vanishing mean oscillation, partially BMO coefficients, Sobolev spaces.

H. Dong was partially supported by NSF grant number DMS-0635607 from IAS, and NSF grant number DMS-0800129.

There is a vast literature of the L_p theory of second-order parabolic and elliptic equations with discontinuous coefficients. It is of particular interest because of its various important applications in nonlinear equations and its subtle links with the theory of stochastic processes. For instance, one implication of the L_p -solvability of parabolic equations in nondivergence form is the weak uniqueness of solutions to associated Itô equations.

For equations with uniformly continuous leading coefficients, the solvability is classical. The L_p theory of second-order equations with discontinuous coefficients was studied extensively in the last three decades. One important class of discontinuous coefficients contains functions with vanishing mean oscillation (VMO), the study of which was initiated in [4] about twenty years ago and continued in [5] and [2]; see also [26, 27, 13] and the references in [22]. The proofs in [4, 5, 2] are based on the Calderón–Zygmund estimate and the Coifman–Rochberg–Weiss commutator estimate. Before that, the L_p theory had been established for some other types of discontinuous coefficients; see, for instance, [25] and [6].

In [21], Krylov gave a unified approach to investigating the L_p -solvability of both divergence and nondivergence form parabolic and elliptic equations with a^{ij} VMO in the spatial variables (and measurable in the time variable in the parabolic case). The proofs in [21] rely mainly on pointwise estimates of sharp functions of spatial derivatives of solutions together with the Stein–Fefferman theorem and the Hardy–Littlewood theorem. By doing this, VMO coefficients are treated in a rather straightforward manner. This result was later improved and generalized in a series of papers [7]–[11], [14]–[18], and [22]–[23]*. For other results about equations and systems with VMO/BMO coefficients, we refer the reader to [1, 3] and references therein.

In contrast, the theory of elliptic and parabolic equations with partially VMO coefficients is quite new, and was originated in [17]. In [17] the W_p^2 -solvability for any $p > 2$ was established for nondivergence form elliptic equations with leading coefficients measurable in one variable and VMO in the others. This result was extended in [18] to parabolic equations. For nondivergence form parabolic equations, further extensions were made later in [14, 15, 16], in which most leading coefficients are measurable in the time variable as well as one spatial variable, and VMO, or BMO with small seminorms, in the other variables. In all these papers, except for some special cases, it is always assumed that $p > 2$. The corresponding results for divergence form equations were proved more recently in [7, 9, 10] by using the idea of “breaking the symmetry” of the coordinates and applying a generalized Stein–Fefferman theorem proved in [23]. In these papers, the solvability is obtained for any $p \in (1, \infty)$ thanks to a standard duality argument for divergence form equations.

*Although the results in some of these papers are claimed for equations with VMO coefficients, the proofs there only require a^{ij} to have locally small mean oscillations.

Roughly speaking, the restriction $p > 2$ mentioned above is due to the following reason. In [17], for instance, a sharp function estimate is deduced from the W_2^2 -solvability of equations with a^{ij} depending only on x^1 , which is obtained by using the method of Fourier transforms. In turn, the right-hand side of the estimate contains maximal functions of q -th power of D^2u for some $q > 2$, which can be made arbitrarily close to 2. Therefore, to apply the Fefferman–Stein theorem and the Hardy–Littlewood maximal function theorem one requires $p \geq q > 2$.

Thus, a natural question is:

Do we still have the L_p -solvability for nondivergence equations with a^{ij} in the class of [14]–[17] when p gets below 2?

In other word, we would like to understand whether the condition $p > 2$ in these papers is necessary, or merely due to technical reasons in the method used. This is one of the motivations of this article. A partial answer to this question can be found in Kunstmann [19], in which the author extended the result in [18] from $p \in (2, \infty)$ to $p \in (p_0, \infty)$ for some $p_0 \in (1, 2)$ in the special case that a^{ij} are time-independent.

In this article, we shall give an affirmative answer to this question in full. One of our main results is the $W_p^{1,2}$ -solvability of nondivergence form parabolic equations for any $p \in (1, \infty)$ (stated in Theorems 2.2 and 2.3), under the assumptions that a^{ij} are partially BMO. More precisely, a^{ij} are assumed to be measurable in one spatial direction and the time variable, and have locally small mean oscillations in the other variables; additionally, we assume a^{11} to be measurable in one spatial direction (or the time variable) and have locally small mean oscillations in the others; see Section 2 for the definitions. Consequently, we derive the W_p^2 -solvability of nondivergence elliptic equations for any $p \in (1, \infty)$, when a^{ij} are in the class of [17]. We also obtain the solvability for nondivergence form parabolic and elliptic *systems* with partially BMO coefficients, which, to the best of our knowledge, is new even when $p > 2$; see Remark 5.3. Moreover, the method used in this paper allows us to treat more general coefficients, i.e. “hierarchically” partially BMO coefficients; see Section 6.

Now we give a brief description of our methods. In order to get below 2, we first establish for any $p \in (1, \infty)$ the $W_p^{1,2}$ -solvability of parabolic equations with a^{11} depending only on either t or x^1 , and $a^{ij}, ij > 1$ depending only on (t, x^1) as stated in Theorems 3.1 and 4.1. To this end, our idea is that in this situation the equation can be rewritten into a divergence form after a suitable change of variables, so that some results recently proved in [7] and [10] for divergence form equations can be applied; see Appendix. In some sense, this idea is reminiscent of the classical argument of deriving from the De Giorgi–Moser–Nash estimate the $C^{1,\alpha}$ regularity for nondivergence elliptic equations with *measurable* coefficients on the plane; see, for instance, [12, §11.2]. Secondly, we bound D^2u and u_t by a portion of D^2u and Pu . Finally, we get a sharp function estimate of the portion of

D^2u by using the Krylov–Safonov interior Hölder estimate. The above three steps are combined together to prove Theorems 2.2 and 2.3 by applying the Fefferman–Stein theorem and the Hardy–Littlewood maximal function theorem. Of course, the Krylov–Safonov estimate is not applicable to systems, for which case we use a bootstrap argument to get an interior Hölder estimate. To deal with equations with hierarchically partially BMO coefficients, in Section 6 we first need to establish the corresponding solvability results for divergence form equations with the leading coefficients in the same class. This is achieved by exploiting the aforementioned idea in [7, 9, 10] of “breaking the symmetry” of the coordinates and running an induction argument on the number of the coordinates.

We note that the $W_p^{1,2}$ -solvability results in this paper admit an extension to the mixed norm spaces $W_{q,p}^{1,2}$ by following the idea in [22]; see also, for instance, [16] and [7]. Here we do not pursue this, and leave it to the interested reader.

A brief outline of the paper: in the next section, we introduce the notation and state some main results, Theorems 2.2, 2.3 and 2.5. Section 3 and 4 are devoted to the proof of Theorems 2.2 and 2.3. We give a generalization of Theorems 2.2 to systems in Section 5. Finally, in Section 6 we discuss equations with hierarchically partially BMO coefficients.

2. NOTATION AND SOME MAIN RESULTS

2.1. Notation. Let $d \geq 1$ be an integer. A typical point in \mathbb{R}^d is denoted by $x = (x^1, \dots, x^d) = (\bar{x}^1, x')$. We set

$$D_i u = u_{x^i}, \quad D_{ij} u = u_{x^i x^j}, \quad D_t u = u_t.$$

By Du and D^2u we mean the gradient and the Hessian matrix of u . On many occasions we need to take these objects relative to only part of variables. We also use the following notation:

$$D_{x'} u = u_{x'}, \quad D_{x^1 x'} u = u_{x^1 x'}, \quad D_{xx'} u = u_{xx'}.$$

Throughout the paper, we always assume that $1 < p, q < \infty$ unless explicitly specified otherwise. By $N(d, p, \dots)$ we mean that N is a constant depending only on the prescribed quantities d, p, l, \dots . For a function $f(t, x)$ in \mathbb{R}^{d+1} , we set

$$(f)_{\mathcal{D}} = \frac{1}{|\mathcal{D}|} \int_{\mathcal{D}} f(t, x) dx dt = \oint_{\mathcal{D}} f(t, x) dx dt,$$

where \mathcal{D} is an open subset in \mathbb{R}^{d+1} and $|\mathcal{D}|$ is the $d+1$ -dimensional Lebesgue measure of \mathcal{D} . For $-\infty \leq S < T \leq \infty$, we denote

$$\begin{aligned} W_p^{1,2}((S, T) \times \mathbb{R}^d) &= \left\{ u : u, u_t, Du, D^2u \in L_p((S, T) \times \mathbb{R}^d) \right\}, \\ \mathcal{H}_p^1((S, T) \times \mathbb{R}^d) &= (1 - \Delta)^{1/2} W_p^{1,2}((S, T) \times \mathbb{R}^d), \\ \mathbb{H}_p^{-1}((S, T) \times \mathbb{R}^d) &= (1 - \Delta)^{1/2} L_p((S, T) \times \mathbb{R}^d). \end{aligned}$$

We also use the abbreviations $L_p = L_p(\mathbb{R}^{d+1})$, $\mathcal{H}_p^1 = \mathcal{H}_p^1(\mathbb{R}^{d+1})$, etc. For any $T \in (-\infty, \infty]$, we denote

$$\mathbb{R}_T = (-\infty, T), \quad \mathbb{R}_T^{d+1} = \mathbb{R}_T \times \mathbb{R}^d.$$

For any integer $k \geq 1$ and $x \in \mathbb{R}^k$, we denote $B_r^k(x)$ to be the k -dimensional cube

$$\{y \in \mathbb{R}^k \mid \max_i |y^i - x^i| < r\}.$$

Set

$$Q_r^k(t, x) = (t - r^2, t) \times B_r^k(x), \quad B_r^k = B_r^k(0), \quad Q_r^k = Q_r^k(0, 0).$$

In case $k = d$ or $d - 1$, we use the abbreviations

$$\begin{aligned} B_r(x) &= B_r^d(x), \quad Q_r(t, x) = Q_r^d(t, x), \\ B_r'(x') &= B_r^{d-1}(x'), \quad Q_r'(t, x') = Q_r^{d-1}(t, x'). \end{aligned}$$

Let

$$\mathcal{Q} = \left\{ Q_r(t, x) : (t, x) \in \mathbb{R}^{d+1}, r \in (0, \infty) \right\}.$$

For a function g defined on \mathbb{R}^{d+1} , we denote its (parabolic) maximal and sharp function, respectively, by

$$\begin{aligned} \mathbb{M}g(t, x) &= \sup_{Q \in \mathcal{Q}: (t, x) \in Q} \int_Q |g(s, y)| dy ds, \\ g^\#(t, x) &= \sup_{Q \in \mathcal{Q}: (t, x) \in Q} \int_Q |g(s, y) - (g)_Q| dy ds. \end{aligned}$$

2.2. Main results without the hierarchical structure. To illustrate the main idea, first we deal with the coefficients considered before in [14]–[18]. We assume that $a^{ij}, ij > 1$ are measurable in x^1 and t , and have locally small mean oscillations in the other variables. In addition, we assume that a^{11} are measurable in t (or x^1) and have locally small mean oscillations in the others. To state the assumptions on a^{ij} precisely, for $R > 0$, we denote

$$\begin{aligned} a_{R,1}^{11,\#} &= \sup_{(t_0, x_0) \in \mathbb{R}^{d+1}} \sup_{r \leq R} \int_{Q_r(t_0, x_0)} |a^{11}(t, x) - \bar{a}^{11}(t)| dx dt, \\ a_{R,2}^{11,\#} &= \sup_{(t_0, x_0) \in \mathbb{R}^{d+1}} \sup_{r \leq R} \int_{Q_r(t_0, x_0)} |a^{11}(t, x) - \hat{a}^{11}(x^1)| dx dt, \\ a_R^\# &= \sup_{(t_0, x_0) \in \mathbb{R}^{d+1}} \sup_{r \leq R} \sup_{(i,j) \neq (1,1)} \int_{Q_r(t_0, x_0)} |a^{ij}(t, x) - \bar{a}^{ij}(t, x^1)| dx dt, \end{aligned}$$

where for each $Q_r(t_0, x_0)$,

$$\begin{aligned}\bar{a}^{11}(t) &= \int_{B_r(x_0)} a^{11}(t, x) dx, \\ \hat{a}^{11}(x^1) &= \int_{Q'_r(t_0, x'_0)} a^{11}(s, x^1, y') dy' ds, \\ \bar{a}^{ij}(t, x^1) &= \int_{B'_r(x'_0)} a^{ij}(t, x^1, y') dy', \quad (i, j) \neq (1, 1).\end{aligned}$$

We shall impose part of the following assumptions on the leading coefficients.

Assumption 2.1 (γ).

- i) There exists a positive constant R_0 such that $a_{R_0,1}^{11,\#} \leq \gamma$.
- ii) There exists a positive constant R_0 such that $a_{R_0,2}^{11,\#} \leq \gamma$.
- iii) There exists a positive constant R_0 such that $a_{R_0}^\# \leq \gamma$.

Next we state a few results of the paper.

Theorem 2.2. *For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 2.1 (γ) i) and iii) for any $T \in (-\infty, +\infty]$ the following holds.*

i) For any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$,

$$\begin{aligned}\lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2u\|_{L_p(\mathbb{R}_T^{d+1})} + \|u_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|Pu - \lambda u\|_{L_p(\mathbb{R}_T^{d+1})},\end{aligned}\tag{2.1}$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, δ, p, K , and R_0 , and N depends only on d, δ and p .

ii) For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique solution $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ of equation (1.1) in \mathbb{R}_T^{d+1} .

iii) In the case that $a^{11} = a^{11}(t)$, $a^{jk} = a^{jk}(t, x^1)$, $jk > 1$ and $b^j \equiv c \equiv 0$, we can take $\lambda_0 = 0$ in i) and ii).

In Section 5, we will extend Theorem 2.2 to parabolic systems under the same regularity assumption on the leading coefficients.

Theorem 2.3. *For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 2.1 (γ) ii) and iii) for any $T \in (-\infty, +\infty]$ the following holds.*

i) For any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$, we have (2.1) provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, δ, p, K , and R_0 , and N depends only on d, δ and p .

ii) For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique solution $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ of equation (1.1) in \mathbb{R}_T^{d+1} .

iii) In the case that $a^{11} = a^{11}(x^1)$, $a^{jk} = a^{jk}(t, x^1)$, $jk > 1$ and $b^j \equiv c \equiv 0$, we can take $\lambda_0 = 0$ in i) and ii).

As a consequence of Theorem 2.3, we obtain the W_p^2 -solvability of second-order elliptic equations in nondivergence form with partially BMO coefficients with locally small semi-norms. For this purpose, we assume that a^{ij} , b^i , c and f are independent of t . For $R > 0$ now we denote

$$a_R^\# = \sup_{x_0 \in \mathbb{R}^d} \sup_{r \leq R} \sup_{i,j} \int_{B_r(x_0)} |a^{ij}(x) - \bar{a}^{ij}(x^1)| dx,$$

where for each $B_r(x_0)$ and (i, j) ,

$$\bar{a}^{ij}(x^1) = \int_{B'_r(x'_0)} a^{ij}(x^1, y') dy'.$$

Assumption 2.4 (γ). There exists a positive constant R_0 such that $a_{R_0}^\# \leq \gamma$.

Theorem 2.5. *For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 2.4 (γ) the following holds.*

i) *For any $u \in W_p^2(\mathbb{R}^d)$,*

$$\lambda \|u\|_{L_p(\mathbb{R}^d)} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}^d)} + \|D^2u\|_{L_p(\mathbb{R}^d)} \leq N \|Lu - \lambda u\|_{L_p(\mathbb{R}^d)},$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depend only on d, δ, p, K and R_0 , and N depends only on d, δ and p .

ii) *For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}^d)$, there exists a unique solution $u \in W_p^2(\mathbb{R}^d)$ of equation (1.2) in \mathbb{R}^d .*

iii) *In the case that $a^{jk} = a^{jk}(x^1)$ and $b^j \equiv c \equiv 0$, we can take $\lambda_0 = 0$ in i) and ii).*

In Section 6, we shall extend Theorems 2.2, 2.3 and 2.5 to equations with more general coefficients. Roughly speaking, we allow hierarchically partially BMO coefficients. This is also one of the main advances of this paper. For example, when $d = 3$, Assumption 6.2 means: a^{11} is measurable in x^1 and has small mean oscillation in (t, x^2, x^3) ; a^{12} , a^{21} and a^{22} are measurable in (t, x^1) and has small mean oscillation in (x^2, x^3) ; finally, all the other a^{ij} are measurable in (t, x^1, x^2) and has small mean oscillation in x^3 .

3. PROOF OF THEOREM 2.2

We prove Theorem 2.2 in this section. Let

$$P_0 u = -u_t + a^{ij} D_{ij} u,$$

where $a^{11} = a^{11}(t)$ and $a^{ij} = a^{ij}(t, x^1)$ for $ij > 1$.

Our proof relies mainly on the following solvability result.

Theorem 3.1. *Let $p \in (1, \infty)$ and $T \in (-\infty, \infty]$. Then for any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ and $\lambda \geq 0$, we have*

$$\begin{aligned} \lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2u\|_{L_p(\mathbb{R}_T^{d+1})} + \|u_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|P_0u - \lambda u\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned} \quad (3.1)$$

where $N = N(d, p, \delta) > 0$. Moreover, for any $f \in L_p(\mathbb{R}_T^{d+1})$ and $\lambda > 0$ there is a unique $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ solving

$$P_0u - \lambda u = f \quad \text{in } \mathbb{R}_T^{d+1}.$$

Proof. First we assume $T = \infty$. By the method of continuity, it suffices to prove the a priori estimate (3.1). Also we can assume $u \in C_0^\infty$ by a density argument. Let $f = P_0u - \lambda u$. Note that

$$-u_t + a^{11}D_1^2u + \Delta_{d-1}u - \lambda u = f + \sum_{ij>1} (\delta_{ij} - a^{ij})D_{ij}u. \quad (3.2)$$

By using Theorem 5.5 of [22], the left-hand side of (3.1) is less than

$$N\|f\|_{L_p} + N \sum_{ij>1} \|D_{ij}u\|_{L_p}.$$

Thus, we only need to prove

$$\|D_{xx'}u\|_{L_p} \leq N\|f\|_{L_p}. \quad (3.3)$$

Rewrite operator P_0 into divergence form

$$P_0u = -u_t + D_1(a^{11}D_1u) + \sum_{j=2}^d D_j((a^{1j} + a^{j1})D_1u) + \sum_{i,j=2}^d D_j(a^{ij}D_iu).$$

Clearly, for $k = 2, \dots, d$, D_ku satisfies

$$P_0(D_ku) - \lambda D_ku = D_kf.$$

Owing to Proposition A.3, we have the \mathcal{H}_p^1 -estimate of divergence form parabolic equations with leading coefficients under the condition of the theorem (and a^{ij} not necessary to be symmetric). Therefore, we get

$$\|D_{xx^k}u\|_{L_p} \leq N\|f\|_{L_p}.$$

This gives (3.3) and completes the proof of the theorem when $T = \infty$.

For general $T \in (-\infty, \infty]$, we use the fact that $u = w$ for $t < T$, where $w \in W_p^{1,2}$ solves

$$P_0w - \lambda w = \chi_{t<T}(P_0u - \lambda u).$$

The theorem is proved. \square

An immediate corollary of Theorem 3.1 is the following solvability result of the initial value problem.

Corollary 3.2. *Let $p \in (1, \infty)$ and $T \in (0, \infty)$. Then for any $f \in L_p((0, T) \times \mathbb{R}^d)$ and $\lambda > 0$ there is a unique $u \in W_p^{1,2}((0, T) \times \mathbb{R}^d)$ solving $P_0 u - \lambda u = f$ in $(0, T) \times \mathbb{R}^d$ and $u(0, \cdot) = 0$. Moreover, we have*

$$\begin{aligned} & \lambda \|u\|_{L_p((0,T) \times \mathbb{R}^d)} + \sqrt{\lambda} \|Du\|_{L_p((0,T) \times \mathbb{R}^d)} + \|D^2 u\|_{L_p((0,T) \times \mathbb{R}^d)} \\ & + \|u_t\|_{L_p((0,T) \times \mathbb{R}^d)} \leq N \|P_0 u - \lambda u\|_{L_p((0,T) \times \mathbb{R}^d)}, \end{aligned}$$

where $N = N(d, p, \delta) > 0$.

To prove the solvability for general operator P , we need the following Hölder estimate, which is an immediate consequence of the Krylov–Safonov estimate.

Lemma 3.3. *Let $p \in (1, \infty)$, $d \geq 2$, $\kappa \geq 2$, and $r > 0$. Assume that $u \in C_0^\infty$ and $P_0 u = 0$ in $Q_{\kappa r}$. Then there exist constants $N = N(d, p, \delta)$ and $\alpha = \alpha(d, p, \delta) \in (0, 1]$ such that for any multi-index $\gamma = (\gamma^1, \gamma')$,*

$$\int_{Q_r} |D^{\gamma'} u - (D^{\gamma'} u)_{Q_r}|^p dx dt \leq N \kappa^{-p\alpha} \left(|D^{\gamma'} u|^p \right)_{Q_{\kappa r}}.$$

Proof. Since $LD^{\gamma'} u = 0$ in $Q_{\kappa r}$, it suffices to concentrate on $\gamma = 0$. By using scaling we reduce the general situation to the one in which $r = 1$. By Lemma 4.2.4 of [20] and Theorem 7.21 of [24],

$$\text{osc}_{Q_{1/\kappa}} u \leq N \kappa^{-\alpha} \|u\|_{L_p(Q_1)}$$

with α and N as in the statement. Scaling this estimate shows that

$$\text{osc}_{Q_1} u \leq N \kappa^{-\alpha} (|u|^p)_{Q_\kappa}^{1/p}.$$

It only remains to observe that

$$\int_{Q_1} |u - (u)_{Q_1}|^p dx dt \leq N (\text{osc}_{Q_1} u)^p.$$

The lemma is proved. \square

Theorem 3.4. *Let $d \geq 2$, $p \in (1, \infty)$ and let α be the constant in Lemma 3.3. Then there is a constant N depending only on d, p, δ such that for any $u \in W_{p,loc}^{1,2}$, $r \in (0, \infty)$, and $\kappa \geq 4$,*

$$\left(|D_{x'}^2 u(t, x) - (D_{x'}^2 u)_{Q_r}|^p \right)_{Q_r} \leq N \kappa^{d+2} (|P_0 u|^p)_{Q_{\kappa r}} + N \kappa^{-p\alpha} (|D_{x'}^2 u|^p)_{Q_{\kappa r}}.$$

Proof. The theorem follows from Corollary 3.2 and Lemma 3.3; see, for instance, the proof of Theorem 4.5 [11]. \square

Next we generalize Theorem 3.4.

Theorem 3.5. *Let $d \geq 2$, $p \in (1, \infty)$ and let α be the constant in Lemma 3.3, $\gamma > 0$, $\tau, \sigma \in (1, \infty)$, $1/\tau + 1/\sigma = 1$. Assume $b^i = c = 0$ and $u \in W_p^{1,2}$. Then under Assumption 2.1 (i) and (iii) there exists a positive constant*

N depending only on d, p, δ , and τ such that, for any $(t_0, x_0) \in \mathbb{R}^{d+1}$, $r \in (0, \infty)$, and $\kappa \geq 4$,

$$\begin{aligned} & (|D_{x'}^2 u(t, x) - (D_{x'}^2 u)_{Q_r(t_0, x_0)}|^p)_{Q_r(t_0, x_0)} \\ & \leq N\kappa^{d+2} (|Pu|^p)_{Q_{\kappa r}(t_0, x_0)} + N\kappa^{d+2} \gamma^{1/\sigma} (|D^2 u|^{p\tau})_{Q_{\kappa r}(t_0, x_0)}^{1/\tau} \\ & \quad + N\kappa^{-p\alpha} (|D_{x'}^2 u|^p)_{Q_{\kappa r}(t_0, x_0)}, \end{aligned} \quad (3.4)$$

provided that u vanishes outside Q_{R_0} .

Proof. Set $f = Pu$. We fix $(t_0, x_0) \in \mathbb{R}^{d+1}$, $\kappa \geq 4$, and $r \in (0, \infty)$. Choose Q to be $Q_{\kappa r}(t_0, x_0)$ if $\kappa r < R_0$ and Q_{R_0} if $\kappa r \geq R_0$. Recall the definitions of $\bar{a}^{11}(t)$ and $\bar{a}^{ij}(t)$ given before Assumption 2.1 i) and iii), respectively. We set

$$\hat{f} = -u_t + \bar{a}^{ij} D_{ij} u.$$

by Theorem 3.4 with an appropriate translation and \bar{a} in place of a ,

$$\begin{aligned} & \int_{Q_r(t_0, x_0)} |D_{x'}^2 u - (D_{x'}^2 u)_{Q_r(t_0, x_0)}|^p dx dt \\ & \leq N\kappa^{d+2} (|\hat{f}|^p)_{Q_{\kappa r}(t_0, x_0)} + N\kappa^{-p\alpha} (|D_{x'}^2 u|^p)_{Q_{\kappa r}(t_0, x_0)}, \end{aligned} \quad (3.5)$$

where N depends only on d, p and δ . By the definition of \hat{f} ,

$$\int_{Q_{\kappa r}(t_0, x_0)} |\hat{f}|^p dx dt \leq N \int_{Q_{\kappa r}(t_0, x_0)} |f|^p dx dt + NI, \quad (3.6)$$

where

$$\begin{aligned} I &= \int_{Q_{\kappa r}(t_0, x_0)} |(\bar{a}^{ij} - a^{ij}) D_{ij} u|^p dx dt \\ &= \int_{Q_{\kappa r}(t_0, x_0) \cap Q_{R_0}} |(\bar{a}^{ij} - a^{ij}) D_{ij} u|^p dx dt. \end{aligned}$$

By Hölder's inequality, we have

$$I \leq NI_1^{1/\sigma} I_2^{1/\tau}, \quad (3.7)$$

where

$$\begin{aligned} I_1 &= \sum_{i,j} \int_{Q_{\kappa r}(t_0, x_0) \cap Q_{R_0}} |\bar{a}^{ij} - a^{ij}|^{p\sigma} dx dt, \\ I_2 &= \int_{Q_{\kappa r}(t_0, x_0)} |D^2 u|^{p\tau} dx dt. \end{aligned}$$

Due to Assumption 2.1 (γ) i) and iii),

$$I_1 \leq \sum_{i,j} \int_Q |\bar{a}^{ij} - a^{ij}|^{p\sigma} dx dt \leq N\gamma|Q| \leq N(\kappa r)^{d+2}\gamma,$$

where $|Q|$ is the volume of Q . This together with (3.5)-(3.7) yields (3.4). The theorem is proved. \square

Lemma 3.6. *Assume $b^i = c = 0$. For any $p \in (1, \infty)$ there exists a constant $\gamma_0 = \gamma_0(d, \delta, p) > 0$ such that under Assumption 2.1 (γ_0) i) for any $u \in C_0^\infty(Q_{R_0})$, we have*

$$\|D^2u\|_{L_p} + \|u_t\|_{L_p} \leq N\|Pu\|_{L_p} + N\|D_{x'}^2u\|_{L_p}, \quad (3.8)$$

where $N = N(d, \delta, p)$. In particular, in case $d = 1$, we have

$$\|D^2u\|_{L_p} + \|u_t\|_{L_p} \leq N\|Pu\|_{L_p}.$$

Proof. We write

$$-u_t + a^{11}D_1^2u + \Delta_{d-1}u = Pu + \sum_{ij>1}(\delta_{ij} - a^{ij})D_{ij}u.$$

The coefficients on the left-hand side above was studied in [21]. By Corollary 3.7 [21], it holds that

$$\|D^2u\|_{L_p} + \|u_t\|_{L_p} \leq N\|Pu\|_{L_p} + N\|D_{xx'}u\|_{L_p}.$$

Finally, to conclude (3.8) it suffices to notice that for any $\varepsilon > 0$,

$$\|D_{x^1x'}u\|_{L_p} \leq \varepsilon\|D_1^2u\|_{L_p} + N(d, p)\varepsilon^{-1}\|D_{x'}^2u\|_{L_p}, \quad (3.9)$$

which is deduced from

$$\|D_{x^1x'}u\|_{L_p} \leq N\|\Delta u\|_{L_p} \leq N\|D_1^2u\|_{L_p} + N\|D_{x'}^2u\|_{L_p}$$

by scaling in x^1 . \square

Lemma 3.7. *Let $p \in (1, \infty)$ and $f \in L_p$. Assume $b^i = c = 0$. Then there exist positive constants γ and N depending only on d, p and δ such that under Assumption 2.1 (γ) i) and iii), for any $u \in W_p^{1,2}$ vanishing outside Q_{R_0} and satisfying $Pu = f$, we have*

$$\|u_t\|_{L_p} + \|D^2u\|_{L_p} \leq N\|f\|_{L_p}.$$

Proof. The case $d = 1$ follows from Lemma 3.6. In the sequel, we assume $d \geq 2$. We fix two numbers $q = (1 + p)/2 \in (1, p)$, $\tau = (1 + 3p)/(2 + 2p) \in (1, \infty)$ so that $p > \tau q$. Let α and γ_0 be the constants in Lemma 3.3 and Lemma 3.6 respectively. Let $\gamma \in (0, \gamma_0)$ be a number to be specified later. Inequality (3.4) with q in place of p implies that on \mathbb{R}^{d+1}

$$\begin{aligned} (D_{x'}^2u)^\# &\leq N\kappa^{(d+2)/q}\mathbb{M}^{1/q}(|f|^q) \\ &\quad + N\kappa^{(d+2)/q}\gamma^{1/(q\sigma)}\mathbb{M}^{1/(q\tau)}(|D^2u|^{q\tau}) + N\kappa^{-\alpha}\mathbb{M}^{1/q}(|D^2u|^q). \end{aligned}$$

We apply the Fefferman–Stein theorem on sharp functions, which reads that for $p \in (1, \infty)$ the L_p norm of a function is bounded by the L_p norm of its sharp function, and the Hardy–Littlewood maximal function theorem, which

reads that the L_p norm of the maximal function is bounded by the L_p norm of the function itself. We then get

$$\begin{aligned} \|D_{x'}^2 u\|_{L_p} &\leq N \|(D_{x'}^2 u)^\# \|_{L_p} \leq N \kappa^{(d+2)/q} \|\mathbb{M}(|f|^q)\|_{L_{p/q}}^{1/q} \\ &\quad + N \kappa^{(d+2)/q} \gamma^{1/(q\sigma)} \|\mathbb{M}(|D^2 u|^{q\tau})\|_{L_{p/(q\tau)}}^{1/(q\tau)} + N \kappa^{-\alpha} \|\mathbb{M}(|D^2 u|^q)\|_{L_{p/q}}^{1/q} \\ &\leq N \kappa^{(d+2)/q} \|f\|_{L_p} + N \left(\kappa^{(d+2)/q} \gamma^{1/(q\sigma)} + \kappa^{-\alpha} \right) \|D^2 u\|_{L_p}, \end{aligned}$$

where in the last inequality we use the fact that $p > \tau q$. From this estimate and Lemma 3.6, we have

$$\|D^2 u\|_{L_p} + \|u_t\|_{L_p} \leq N \kappa^{(d+2)/q} \|f\|_{L_p} + N \left(\kappa^{(d+2)/q} \gamma^{1/(q\sigma)} + \kappa^{-\alpha} \right) \|D^2 u\|_{L_p}.$$

To finish the proof of the lemma, it suffices to choose a large κ and then a small γ so that

$$N \left(\kappa^{(d+2)/q} \gamma^{1/(q\sigma)} + \kappa^{-\alpha} \right) \leq 1/2.$$

□

Now we are ready to prove Theorem 2.2.

Proof of Theorem 2.2. To prove i), as in the proof of Theorem 3.1, we only need to verify (2.1) for $T = \infty$ and $u \in C_0^\infty$. This in turn is obtained from Lemma 3.7 by using a partition of unity and an idea of Agmon (see, for instance, [21]). Assertion ii) is obtained from i) by the method of continuity, and Assertion iii) is already proved in Theorem 3.1. The theorem is proved. □

4. PROOF OF THEOREM 2.3

The objective of this section is to prove Theorem 2.3. We follow the strategy in the previous section. However, the proofs here are more involved.

Let

$$P_0 u = -u_t + a^{ij} D_{ij} u,$$

where

$$a^{11} = a^{11}(x^1), \quad a^{ij} = a^{ij}(t, x^1) \quad \text{for } ij > 1.$$

Similarly we need a solvability result for equations with these simple leading coefficients.

Theorem 4.1. *Let $p \in (1, \infty)$ and $T \in (-\infty, \infty]$. Then for any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ and $\lambda \geq 0$, we have*

$$\begin{aligned} \lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2 u\|_{L_p(\mathbb{R}_T^{d+1})} + \|u_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|P_0 u - \lambda u\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned} \tag{4.1}$$

where $N = N(d, p, \delta) > 0$. Moreover, for any $f \in L_p(\mathbb{R}_T^{d+1})$ and $\lambda > 0$ there is a unique $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ solving

$$P_0 u - \lambda u = f \quad \text{in } \mathbb{R}_T^{d+1}.$$

Proof. As in the proof of Theorem 3.1, it suffices to prove the a priori estimate (4.1) for $u \in C_0^\infty$ and $T = \infty$. Let $\lambda > 0$ and $f = P_0 u - \lambda u$. Here we are not able to directly write P_0 into a divergence form operator. However, this can be done after a change of variables:

$$y^1 = \phi(x^1) := \int_0^{x^1} \frac{1}{a^{11}(s)} ds, \quad y^j = x^j, \quad j \geq 2.$$

It is easy to see that ϕ is a bi-Lipschitz map and

$$\delta \leq y^1/x^1 \leq \delta^{-1}, \quad D_{y^1} = a^{11}(x^1)D_{x^1}.$$

Denote

$$\begin{aligned} v(t, y^1, y') &= u(t, \phi^{-1}(y^1), y'), \\ \bar{a}^{11}(y^1) &= a^{11}(\phi^{-1}(y^1)), \\ \bar{a}^{ij}(t, y^1) &= a^{ij}(t, \phi^{-1}(y^1)), \quad ij > 1, \\ \bar{f}(t, y) &= f(t, \phi^{-1}(y^1), y'). \end{aligned}$$

In the (t, y) -coordinates, define a divergence form operator \bar{P}_0 by

$$\bar{P}_0 v = -v_t + D_1 \left(\frac{1}{\bar{a}^{11}} D_1 v \right) + \sum_{j=2}^d D_j \left(\frac{\bar{a}^{1j} + \bar{a}^{j1}}{\bar{a}^{11}} D_1 v \right) + \sum_{i,j=2}^d D_j (\bar{a}^{ij} D_i v).$$

It is easily seen that \bar{P}_0 is uniformly nondegenerate with an ellipticity constant depending only on δ . A simple calculation shows that v satisfies in \mathbb{R}^{d+1}

$$\bar{P}_0 v - \lambda v = \bar{f}.$$

By Proposition A.1, we have

$$\lambda \|v\|_{L_p} + \sqrt{\lambda} \|Dv\|_{L_p} \leq N \|\bar{f}\|_{L_p}.$$

Therefore,

$$\lambda \|u\|_{L_p} + \sqrt{\lambda} \|Du\|_{L_p} \leq N \|f\|_{L_p}. \quad (4.2)$$

Next we estimate $D^2 u$. Notice that for each $k = 2, \dots, d$ $D_k v$ satisfies

$$\bar{P}_0(D_k v) - \lambda D_k v = D_k \bar{f}.$$

Again by using Proposition A.1, we get

$$\|D_{yy^k} v\|_{L_p} \leq N \|\bar{f}\|_{L_p},$$

which implies

$$\|D_{xx'} u\|_{L_p} \leq N \|f\|_{L_p}. \quad (4.3)$$

Finally, to estimate $D_1^2 u$, we return to the equation in the original coordinate system. From (3.2), we see that $w := D_1 u$ satisfies

$$-w_t + D_1(a^{11} D_1 w) + \Delta_{d-1} w - \lambda w = D_1 f + \sum_{ij>1} D_1 ((\delta_{ij} - a^{ij}) D_{ij} u). \quad (4.4)$$

We use Proposition A.1 again to get

$$\|D_1^2 u\|_{L_p} \leq \|Dw\|_{L_p} \leq N\|f\|_{L_p} + N \sum_{ij>1} \|D_{ij}u\|_{L_p}. \quad (4.5)$$

Combining (4.2), (4.3) and (4.5) yields (4.1) by bearing in mind that

$$u_t = a^{ij}D_{ij}u - \lambda u - f. \quad (4.6)$$

The theorem is proved. \square

Observe that in the proofs of Lemma 3.6, Theorems 3.4 and 3.5, we only used the solvability of parabolic equations with “simple” leading coefficients and the fact that a^{ij} are independent of x' . Since Theorem 4.1 is proved, we still have (3.4) if Assumption 2.1 i) in Theorem 3.5 is replaced by Assumption 2.1 ii). More precisely,

Theorem 4.2. *Let $d \geq 2$, $p \in (1, \infty)$ and let α be the constant in Lemma 3.3, $\gamma > 0$, $\tau, \sigma \in (1, \infty)$, $1/\tau + 1/\sigma = 1$. Assume $b^i = c = 0$ and $u \in W_p^{1,2}$. Then under Assumption 2.1 (γ) ii) and iii) there exists a positive constant N depending only on d , p , δ , and τ such that, for any $(t_0, x_0) \in \mathbb{R}^{d+1}$, $r \in (0, \infty)$, and $\kappa \geq 4$, we have*

$$\begin{aligned} & (|D_{x'}^2 u(t, x) - (D_{x'}^2 u)_{Q_r(t_0, x_0)}|^p)_{Q_r(t_0, x_0)} \leq N\kappa^{d+2} (|Pu|^p)_{Q_{\kappa r}(t_0, x_0)} \\ & + N\kappa^{d+2} \gamma^{1/\sigma} (|D^2 u|^{p\tau})_{Q_{\kappa r}(t_0, x_0)}^{1/\tau} + N\kappa^{-p\alpha} (|D_{x'}^2 u|^p)_{Q_{\kappa r}(t_0, x_0)}, \end{aligned}$$

provided that u vanishes outside Q_{R_0} .

Lemma 4.3. *Assume $b^i = c = 0$. For any $p \in (1, \infty)$ there exists a γ_1 , μ_1 and N , depending only on d , δ and p , such that under Assumption 2.1 (γ_1) ii) for any $u \in C_0^\infty(Q_{\mu_1^{-1}R_0})$, we have*

$$\|D^2 u\|_{L_p} + \|u_t\|_{L_p} \leq N\|Pu\|_{L_p} + N\|D_{x'}^2 u\|_{L_p}. \quad (4.7)$$

Proof. Because $w = D_1 u$ satisfies (4.4), due to Proposition A.2 we have

$$\|D_1^2 u\|_{L_p} \leq \|Dw\|_{L_p} \leq N\|f\|_{L_p} + \sum_{ij>1} \|D_{ij}u\|_{L_p}.$$

This and (4.6) with $\lambda = 0$ yield

$$\|u_t\|_{L_p} + \|D^2 u\|_{L_p} \leq N\|f\|_{L_p} + \sum_{ij>1} \|D_{ij}u\|_{L_p}.$$

To finish the proof of (4.7), it suffices to use (3.9). \square

By using Theorem 4.2 and Lemma 4.3, we can prove Theorem 2.3 in exactly the same way as in the proof of Theorem 2.2. We omit the details.

5. AN EXTENSION TO PARABOLIC SYSTEMS

The objective of this section is to extend Theorem 2.2 to parabolic systems under the same regularity assumption on the leading coefficients. Let $m \geq 2$ be an integer. We consider parabolic operators in nondivergence form

$$P\mathbf{u} = -\mathbf{u}_t + A^{\alpha\beta} D_{\alpha\beta} \mathbf{u} + B^\alpha D_\alpha \mathbf{u} + C\mathbf{u},$$

acting on (column) vector-valued functions $\mathbf{u} = (u^1, \dots, u^m)^T$. Here for each $\alpha, \beta = 1, \dots, d$, $A^{\alpha\beta}$, B^α and C are $m \times m$ matrix-valued functions given on \mathbb{R}^{d+1} ; i.e., $A^{\alpha\beta} = [A_{ij}^{\alpha\beta}(t, x)]_{m \times m}$, etc.

As in the scalar case, all the coefficients are assumed to be bounded and measurable, and $A^{\alpha\beta}$ are uniformly elliptic, i.e.

$$\begin{aligned} |B^\alpha| + |C| &\leq K, \quad |A_{ij}^{\alpha\beta}| \leq \delta^{-1}, \\ \delta \sum_{i=1}^m \sum_{\alpha=1}^d |\xi_\alpha^i|^2 &\leq \sum_{\alpha, \beta=1}^d \sum_{i, j=1}^m A_{ij}^{\alpha\beta}(t, x) \xi_\alpha^i \xi_\beta^j \leq \delta^{-1} \sum_{i=1}^m \sum_{\alpha=1}^d |\xi_\alpha^i|^2 \end{aligned} \quad (5.1)$$

for all $(t, x) \in \mathbb{R}^{d+1}$ and $\xi_\alpha = (\xi_\alpha^i) \in \mathbb{R}^m$, $\alpha = 1, \dots, d$. Additionally, we assume $A^{\beta 1}$ is symmetric matrix for $\beta = 2, \dots, d$.

Similarly we also define the objects \bar{A}^{11} , \bar{A}^{ij} , $ij > 1$, $A_{R,1}^{11,\#}$ and $A_R^\#$ as in Section 2.

Assumption 5.1 (γ). There exists a positive constant R_0 such that

$$A_{R_0,1}^{11,\#} + A_{R_0}^\# \leq \gamma.$$

Denote

$$P_0 \mathbf{u} = -\mathbf{u}_t + A^{\alpha\beta} D_{\alpha\beta} \mathbf{u},$$

where

$$A^{11} = A^{11}(t), \quad A^{\alpha\beta} = A^{\alpha\beta}(t, x^1), \quad \alpha\beta > 1.$$

First we give a counterpart of Theorem 3.1.

Theorem 5.2. *Let $p \in (1, \infty)$ and $T \in (-\infty, \infty]$. Then for any $\mathbf{u} \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ and $\lambda \geq 0$, we have*

$$\begin{aligned} \lambda \|\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|D\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2 \mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \|\mathbf{u}_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|P_0 \mathbf{u} - \lambda \mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned}$$

where $N = N(d, m, p, \delta) > 0$. Moreover, for any $\mathbf{f} \in L_p(\mathbb{R}_T^{d+1})$ and $\lambda > 0$ there is a unique $\mathbf{u} \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ solving

$$P_0 \mathbf{u} - \lambda \mathbf{u} = \mathbf{f} \quad \text{in } \mathbb{R}_T^{d+1}.$$

Proof. The proof is almost the same as that of Theorem 3.1. So we only give a few comments. The only difference is that instead of Theorem 5.5 of [22], we use a system version of it, which can be found in Theorem 3.1 of [8]. Here the symmetric condition of $A^{\beta 1}$, $\beta = 2, \dots, d$ is needed to apply Proposition A.3. \square

Remark 5.3. The results in [18, 14, 15, 16] rely on the $W_2^{1,2}$ -solvability proved in [18] for parabolic equations with a^{ij} depending only on (t, x^1) . The extension of the latter to parabolic systems is unknown to us, except for the one space dimensional case. So is the corresponding extension of W_2^2 -solvability to elliptic systems. This is because the proof in [18] uses the maximum principle, which is only valid for scalar equations.

The key estimate in this section is the following lemma, which is a generalization of Lemma 3.3. We note that, since we are dealing with systems, the Krylov–Safonov estimate does not apply.

Lemma 5.4. *Let $p \in (1, \infty)$, $d \geq 2$, $\kappa \geq 2$, and $r > 0$. Assume that $\mathbf{u} \in C_0^\infty$ and $P_0 \mathbf{u} = 0$ in $Q_{\kappa r}$. Then for any $\alpha \in (0, 1)$, there exist constants $N = N(d, m, p, \delta, \alpha)$ such that for any multi-index $\gamma = (\gamma^1, \gamma')$*

$$\int_{Q_r} |D^{\gamma'} \mathbf{u} - (D^{\gamma'} \mathbf{u})_{Q_r}|^p dx dt \leq N \kappa^{-p\alpha} \left(|D^{\gamma'} \mathbf{u}|^p \right)_{Q_{\kappa r}}. \quad (5.2)$$

Proof. As in the proof of Lemma 3.3, without loss of generality, we may assume $|\gamma| = 0$. We again use the fact that $P_0 \mathbf{u}$ can be rewritten into a divergence form:

$$-\mathbf{u}_t + D_1(A^{11} D_1 \mathbf{u}) + \sum_{\beta=2}^d D_\beta \left((A^{1\beta} + A^{\beta 1}) D_1 \mathbf{u} \right) + \sum_{\alpha, \beta=2}^d D_\beta (A^{\alpha\beta} D_\alpha \mathbf{u}).$$

Now (5.2) follows from Proposition A.4. The lemma is proved. \square

Remark 5.5. We claim a more general estimate: under the assumptions of Lemma 5.4, we have

$$\int_{Q_r} |D^\gamma \mathbf{u} - (D^\gamma \mathbf{u})_{Q_r}|^p dx dt \leq N \kappa^{-p\alpha} \left(|D^{|\gamma^1|} D^{\gamma'} \mathbf{u}|^p \right)_{Q_{\kappa r}}.$$

provided that $\gamma^1 = 0$ or 1 . Indeed, it suffices to consider the case $\gamma^1 = 1$ and $|\gamma'| = 0$. This case can be proved by using the iteration argument in the proof of Lemma 4.2 [10] and the Poincaré's inequality. We leave the details to the interested reader. This estimate will not be used in the sequel.

With Theorem 5.2 and Lemma 5.4 available, we can obtain analogous results of Theorems 3.4 and 3.5. Following the lines of Section 3, we get the following solvability result, which is a generalization of Theorem 2.2.

Theorem 5.6. *For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, m, \delta, p) > 0$ such that under Assumption 5.1 (γ) for any $T \in (-\infty, +\infty]$ the following holds.*

i) *For any $\mathbf{u} \in W_p^{1,2}(\mathbb{R}_T^{d+1})$,*

$$\begin{aligned} \lambda \|\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|D\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2 \mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \|\mathbf{u}_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|P\mathbf{u} - \lambda \mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned}$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, m, δ, p, K and R_0 , and N depends only on d, m, δ and p .

ii) For any $\lambda > \lambda_0$ and $\mathbf{f} \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique solution $\mathbf{u} \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ of

$$P\mathbf{u} - \lambda\mathbf{u} = \mathbf{f} \quad \text{in } \mathbb{R}_T^{d+1}.$$

iii) In the case that $A^{11} = A^{11}(t)$, $A^{\alpha\beta} = A^{\alpha\beta}(t, x^1)$, $\alpha\beta > 1$ and $b^j \equiv c \equiv 0$, we can take $\lambda_0 = 0$ in i) and ii).

We remark that Theorem 5.6 is new even when $p > 2$. It is also worth noting that under Assumption 5.1 (γ) solutions of the parabolic systems are Hölder continuous if $p > (d+2)/2$. Unlike scalar equations, for which we have the Krylov–Safonov Hölder estimate, generally this property is not possessed by solutions to systems (see, for instance, [26]).

6. HIERARCHICALLY PARTIALLY BMO COEFFICIENTS

We consider in this section parabolic equations with more general coefficients. The assumption is that for $ij > 1$ a^{ij} are measurable in $(t, x^1, \dots, x^{\pi_{ij}})$ and BMO in the other coordinates, where

$$\pi_{ij} = \max(i, j) - 1.$$

In addition, we suppose a^{11} is measurable in t (or x^1) and BMO in the other coordinates. More precise assumptions are stated below. To the best of our knowledge, this class of coefficients has been considered before.

We recall the definition of $a_{R,1}^{11,\#}$ and $a_{R,2}^{11,\#}$ in the introduction. For $R > 0$, we denote

$$a_R^\# = \sup_{(t_0, x_0) \in \mathbb{R}^{d+1}} \sup_{r \leq R} \sup_{(i,j) \neq (1,1)} \int_{Q_r(t_0, x_0)} |a^{ij} - \bar{a}^{ij}| dx dt,$$

where for each $Q_r(t_0, x_0)$ and $(i, j) \neq (1, 1)$,

$$\begin{aligned} \bar{a}^{ij} &= \bar{a}^{ij}(t, x^1, \dots, x^{\pi_{ij}}) \\ &= \int_{B_r^{d-\pi_{ij}}(x_0^{\pi_{ij}+1}, \dots, x_0^d)} a^{ij}(t, x^1, \dots, x^{\pi_{ij}}, y^{\pi_{ij}+1}, \dots, y^d) dy^{\pi_{ij}+1} \dots dy^d. \end{aligned}$$

We impose either one of the following assumptions on a^{ij} .

Assumption 6.1 (γ). There exists a positive constant R_0 such that

$$a_{R_0,1}^{11,\#} + a_R^\# \leq \gamma.$$

Assumption 6.2 (γ). There exists a positive constant R_0 such that

$$a_{R_0,2}^{11,\#} + a_R^\# \leq \gamma.$$

Clearly the assumptions above are weaker than those in Theorems 2.2 and 2.3 in terms of the regularity of a^{ij} for $i > 2$ or $j > 2$.

Following the idea in Section 3 and 4, to prove the $W_p^{1,2}$ -solvability of nondivergence form equations under these assumptions, in Subsection 6.1 we first establish the corresponding \mathcal{H}_p^1 -solvability of divergence form equations under the same assumptions on the coefficients. In Subsection 6.2, we prove the solvability for nondivergence equations.

6.1. Solvability of divergence form equations. Let us consider the \mathcal{H}_p^1 -solvability of parabolic equations in divergence form:

$$\mathcal{P}u - \lambda u = \operatorname{div} g + f,$$

where $\lambda \geq 0$ is a constant and

$$\begin{aligned} \mathcal{P}u &= -u_t + D_j(a^{ij}D_i u) + D_i(a^i u) + b^i D_i u + cu, \\ g &= (g^1, g^2, \dots, g^d). \end{aligned}$$

We assume the same boundedness and ellipticity condition on a^{ij} , b^i and c as in the introduction. We also assume that a^i are measurable and bounded by K .

Theorem 6.3. *Let $p \in (1, \infty)$, there exists a constant $\gamma = \gamma(d, p, \delta) > 0$ such that under Assumption 6.1 (γ) (or 6.2 (γ)) for any $T \in (-\infty, +\infty]$ the following holds.*

i) *For any $u \in \mathcal{H}_p^1(\mathbb{R}_T^{d+1})$,*

$$\lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} \leq N(\sqrt{\lambda} + 1) \|\mathcal{P}u - \lambda u\|_{\mathbb{H}_p^{-1}(\mathbb{R}_T^{d+1})}$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, p, δ, K and R_0 , and N depends only on d, p and δ .

ii) *For any $\lambda > \lambda_0$ and $f, g \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique $u \in \mathcal{H}_p^1(\mathbb{R}_T^{d+1})$ solving*

$$\mathcal{P}u - \lambda u = f + \operatorname{div} g$$

in \mathbb{R}_T^{d+1} . And u satisfies the estimate

$$\lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} \leq N\sqrt{\lambda} \|g\|_{L_p(\mathbb{R}_T^{d+1})} + N\|f\|_{L_p(\mathbb{R}_T^{d+1})}.$$

iii) *In the case that $a^{11} = a^{11}(t)$ (or $a^{11} = a^{11}(x^1)$ respectively), and*

$$a^{ij} = a^{ij}(t, x^1, \dots, x^{\pi_{ij}}), \quad ij > 1, \quad a^j = b^j \equiv c \equiv 0,$$

we can take $\lambda_0 = 0$ in i) and ii).

The theorem above generalized the main result of [7]. The idea in the proof is to “break the symmetry” of the coordinates.

For simplicity, we only give a proof of Theorem 6.3 when $d = 3$, or $d > 3$ and π_{ij} is replaced by $\tilde{\pi}_{ij} := \min(\pi_{ij}, 2)$. The general case can be proved by an induction. We remark that the assumptions of Theorems 2.2 and 2.3 (and those in [7], [10]) correspond to the case when π_{ij} is replaced by $\min(\pi_{ij}, 1)$.

First let us consider the situation that Assumption 6.2 holds. We have the following corollary of Proposition A.2.

Corollary 6.4. *Let $p \in (1, \infty)$ and $g \in L_p$. Suppose $a^i \equiv b^i \equiv c \equiv 0$. Let μ_1 and γ_1 be the constants in Proposition A.2. Then there exist constants $\mu_2 \in [\mu_1, \infty)$, $\gamma_2 \in (0, \gamma_1]$ and N depending only on d, p and δ such that,*

under Assumption 6.2 (γ_2) with $\tilde{\pi}_{ij}$ in place of π_{ij} , for any $u \in C_0^\infty(Q_{\mu_2^{-1}R_0})$ satisfying $\mathcal{P}u = \operatorname{div} g$, we have

$$\|Du\|_{L_p} \leq N\|g\|_{L_p} + N \sum_{j=3}^d \|D_j u\|_{L_p}.$$

Proof. Let $\mu_2 = \mu_1 \mu$ and $\gamma_2 = \mu^{-4} \gamma_1$, where $\mu \in [1, \infty)$ is a number to be specified later. Denote

$$\bar{u}(t, x) = u(\mu^{-2}t, \mu^{-1}x^1, \mu^{-1}x^2, x^3, \dots, x^d).$$

It is clear that $\bar{u} \in C_0^\infty(Q_{\mu_1^{-1}R_0})$, and \bar{u} satisfies

$$\begin{aligned} & -\mu^2 \bar{u}_t + \sum_{i,j=1}^2 D_j(\mu^2 \bar{a}^{ij} D_i \bar{u}) + \sum_{i=1}^2 \sum_{j>2} D_j(\mu \bar{a}^{ij} D_i \bar{u})_{x^j} \\ & + \sum_{j=1}^2 \sum_{i>2} D_j(\mu \bar{a}^{ij} D_i \bar{u}) + \sum_{i>2, j>2} D_j(\bar{a}^{ij} D_i \bar{u}) = \operatorname{div}(\bar{g}), \end{aligned}$$

where

$$\begin{aligned} \bar{a}^{ij} &= a^{ij}(\mu^{-2}t, \mu^{-1}x^1, \mu^{-1}x^2, x^3, \dots, x^d), \\ \bar{g}^i &= \mu g^i(\mu^{-2}t, \mu^{-1}x^1, \mu^{-1}x^2, x^3, \dots, x^d), \quad i = 1, 2, \\ \bar{g}^i &= g^i(\mu^{-2}t, \mu^{-1}x^1, \mu^{-1}x^2, x^3, \dots, x^d), \quad i > 2. \end{aligned}$$

Simple calculation shows that \bar{a}^{ij} satisfy Assumption 6.2 (γ_1) with $\tilde{\pi}_{ij}$ in place of π_{ij} .

We define an operator \mathfrak{P} by

$$\mathfrak{P}u = -u_t + \sum_{i,j=1}^2 D_j(\bar{a}^{ij} D_i u) + \sum_{j=3}^d D_j^2 u.$$

Then we have

$$\mathfrak{P}\bar{u} = \operatorname{div} \mathfrak{g},$$

where

$$\begin{aligned} \mathfrak{g}^k &= \mu^{-2} \bar{g}^k - \mu^{-1} \sum_{i>2} \bar{a}^{ik} D_i \bar{u}, \quad k = 1, 2, \\ \mathfrak{g}^k &= \mu^{-2} \bar{g}^k - \mu^{-1} \sum_{i=1}^2 \bar{a}^{ik} D_i \bar{u} - \mu^{-2} \sum_{i>2} \bar{a}^{ik} D_i \bar{u} + D_k \bar{u}, \quad k \geq 3. \end{aligned}$$

Since the coefficients of \mathfrak{P} satisfy Assumption 2.1 (γ_1) ii) and iii), we get from Proposition A.2 that

$$\|D\bar{u}\|_{L_p} \leq N_1 \|\mathfrak{g}\|_{L_p},$$

which implies

$$\sum_{i=1}^2 \|D_i u\|_{L_p} \leq N_1 \|g\|_{L_p} + N_1 \mu^{-1} \sum_{i=1}^2 \|D_i u\|_{L_p} + N_1 \mu \sum_{i=3}^d \|D_i u\|_{L_p},$$

where $N_1 = N_1(d, p, \delta) > 0$. To finish the proof, it suffices to choose $\mu = (2N_1)^{-1}$. \square

Because of Corollary 6.4, we only need to obtain a good estimate of $\|D_j u\|_{L_p}$ for $j \geq 3$. Following the lines of Section 5 [7], we have the next estimate of mean oscillations.

Lemma 6.5. *Let $a^i = b^i = 0$, $c = 0$, $\sigma, \tau \in (1, \infty)$ satisfying $1/\sigma + 1/\tau = 1$. Assume $u \in \mathcal{H}_{2,loc}^1$ and $\mathcal{P}u = \operatorname{div} g$, where $g \in L_{2,loc}$. Then under Assumption 6.2 (γ) with $\tilde{\pi}_{ij}$ in place of π_{ij} , there exist an $\alpha = \alpha(d, \delta) \in (0, 1)$ and a positive constant N depending only on d, σ and δ such that*

$$\begin{aligned} \sum_{i=3}^d (|D_i u - (D_i u)_{Q_r(t_0, x_0)}|^2)_{Q_r(t_0, x_0)} &\leq N\nu^{-2\alpha} (|Du|^2)_{Q_{\nu r}(t_0, x_0)} \\ &\quad + N\nu^{d+2} \left((|g|^2)_{Q_{\nu r}(t_0, x_0)} + \gamma^{1/\sigma} (|Du|^{2\tau})_{Q_{\nu r}(t_0, x_0)}^{1/\tau} \right), \end{aligned}$$

for any $r \in (0, \infty)$, $\nu \geq 4$ and $(t_0, x_0) \in \mathbb{R}^{d+1}$, provided that u vanishes outside Q_{R_0} .

Theorem 6.6. *Let $p > 2$, $a^i = b^i = c = 0$. Let μ_2 and γ_2 be the constants in Corollary 6.4. Then there exists a constant $\gamma_3 \in (0, \gamma_2]$ depending only on d, p and δ such that under Assumption 6.2 (γ_3) with $\tilde{\pi}_{ij}$ in place of π_{ij} , for any $u \in C_0^\infty(Q_{\mu_2^{-1}R_0})$ and $g \in L_p$ satisfying $\mathcal{P}u = \operatorname{div} g$, we have*

$$\|Du\|_{L_p} \leq N\|g\|_{L_p}.$$

Proof. Fix $\nu \geq 4$ and $\gamma_3 \in (0, \gamma_2]$ to be specified later. Let α be the constant in Lemma 6.5 and choose $\tau > 1$ such that $p > 2\tau$. Lemma 6.5 implies

$$\begin{aligned} \sum_{i=3}^d (D_i u)^\# &\leq N\nu^{-\alpha} (\mathbb{M}(|Du|^2))^{1/2} + N\nu^{(d+2)/2} (\mathbb{M}(|g|^2))^{1/2} \\ &\quad + N\nu^{(d+2)/2} \gamma_3^{1/(2\sigma)} (\mathbb{M}(|Du|^{2\tau}))^{1/(2\tau)}. \end{aligned}$$

By using the Fefferman–Stein theorem on sharp functions, and the Hardy–Littlewood maximal function theorem, we get

$$\sum_{i=3}^d \|D_i u\|_{L_p} \leq N(\nu^{-\alpha} + \nu^{(d+2)/2} \gamma_3^{1/(2\sigma)}) \|Du\|_{L_p} + N\nu^{(d+2)/2} \|g\|_{L_p}. \quad (6.1)$$

Since $\gamma_3 \in (0, \gamma_2]$ and $u \in C_0^\infty(Q_{\mu_2^{-1}R_0})$, estimate (6.1) and Corollary 6.4 give

$$\|Du\|_{L_p} \leq N(\nu^{-\alpha} + \nu^{(d+2)/2} \gamma_3^{1/(2\sigma)}) \|Du\|_{L_p} + N\nu^{(d+2)/2} \|g\|_{L_p}. \quad (6.2)$$

We take ν sufficiently large, and γ_3 small such that in (6.2)

$$N(\nu^{-\alpha} + \nu^{(d+2)/2} \gamma_3^{1/(2\sigma)}) \leq 1/2.$$

The theorem is proved. \square

We are now in the position to prove Theorem 6.3.

Proof of Theorem 6.3. Recall that for simplicity we replace π_{ij} in the assumption by $\tilde{\pi}_{ij}$. The general case can be done by an induction.

The case $p = 2$ is classical. Suppose that Assumption 6.2 is true. For $p > 2$, the result follows from Theorem 6.6 by using a partition of unity and an idea by S. Agmon; see, for instance, [21]. The case $1 < p < 2$ follows from a standard duality argument.

For the remaining case that Assumption 6.1 is true, we follow the argument in this subsection and use the results in [10] instead of [7]. The theorem is proved. \square

As before, we have the following solvability result for elliptic equations in divergence form:

$$\mathcal{L}u - \lambda u = \operatorname{div} g + f,$$

where all the coefficients are time-independent and

$$\mathcal{L}u = D_j(a^{ij}D_i u) + D_i(a^i u) + b^i D_i u + cu.$$

Theorem 6.7. *Let $p \in (1, \infty)$, there exists a constant $\gamma = \gamma(d, p, \delta) > 0$ such that under Assumption 6.2 (γ) the following holds.*

i) *For any $u \in W_p^1(\mathbb{R}^d)$,*

$$\lambda \|u\|_{L_p(\mathbb{R}^d)} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}^d)} \leq N(\sqrt{\lambda} + 1) \|\mathcal{L}u - \lambda u\|_{\mathbb{H}_p^{-1}(\mathbb{R}^d)}$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, p, δ, K and R_0 , and N depends only on d, p and δ .

ii) *For any $\lambda > \lambda_0$ and $f, g \in L_p(\mathbb{R}^d)$, there exists a unique $u \in W_p^1(\mathbb{R}^d)$ solving*

$$\mathcal{L}u - \lambda u = f + \operatorname{div} g$$

in \mathbb{R}^d . And u satisfies the estimate

$$\lambda \|u\|_{L_p(\mathbb{R}^d)} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}^d)} \leq N\sqrt{\lambda} \|g\|_{L_p(\mathbb{R}^d)} + N\|f\|_{L_p(\mathbb{R}^d)}.$$

iii) *In the case that*

$$a^{11} = a^{11}(x^1), \quad a^{ij} = a^{ij}(x^1, \dots, x^{\pi_{ij}}), \quad ij > 1, \quad a^j = b^j \equiv c \equiv 0,$$

we can take $\lambda_0 = 0$ in i) and ii).

6.2. Solvability of nondivergence form equations. We prove in this subsection the following solvability theorems, which generalize Theorems 2.2 and 2.3 respectively.

Theorem 6.8. *For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 6.1 (γ) for any $T \in (-\infty, +\infty]$ the following holds.*

i) *For any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$,*

$$\begin{aligned} \lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2 u\|_{L_p(\mathbb{R}_T^{d+1})} + \|u_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|Pu - \lambda u\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned} \tag{6.3}$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, δ, p, K and R_0 , and N depends only on d, δ and p .

ii) For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique solution $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ of equation (1.1) in \mathbb{R}_T^{d+1} .

iii) In the case that $a^{11} = a^{11}(t)$, and

$$a^{ij} = a^{ij}(t, x^1, \dots, x^{\pi_{ij}}), \quad ij > 1, \quad b^j \equiv c \equiv 0,$$

we can take $\lambda_0 = 0$ in i) and ii).

Theorem 6.9. For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 6.2 (γ) for any $T \in (-\infty, +\infty]$ the following holds.

i) For any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$, we have (6.3) provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, δ, p, K and R_0 , and N depends only on d, δ and p .

ii) For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}_T^{d+1})$, there exists a unique solution $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ of equation (1.1) in \mathbb{R}_T^{d+1} .

iii) In the case that $a^{11} = a^{11}(x^1)$, and

$$a^{ij} = a^{ij}(t, x^1, \dots, x^{\pi_{ij}}), \quad ij > 1, \quad b^j \equiv c \equiv 0,$$

we can take $\lambda_0 = 0$ in i) and ii).

As a consequence of Theorem 6.9, we have the following solvability theorem for elliptic equations.

Theorem 6.10. For any $p \in (1, \infty)$, there exists a $\gamma = \gamma(d, \delta, p) > 0$ such that under Assumption 6.2 (γ) the following holds.

i) For any $u \in W_p^2(\mathbb{R}^d)$, we have

$$\lambda \|u\|_{L_p(\mathbb{R}^d)} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}^d)} + \|D^2u\|_{L_p(\mathbb{R}^d)} \leq N \|Lu - \lambda u\|_{L_p(\mathbb{R}^d)},$$

provided that $\lambda \geq \lambda_0$, where $\lambda_0 \geq 0$ depends only on d, δ, p, K and R_0 , and N depends only on d, δ and p .

ii) For any $\lambda > \lambda_0$ and $f \in L_p(\mathbb{R}^d)$, there exists a unique solution $u \in W_p^2(\mathbb{R}^d)$ of equation (1.2) in \mathbb{R}^d .

iii) In the case that $a^{11} = a^{11}(x^1)$, and

$$a^{ij} = a^{ij}(x^1, \dots, x^{\pi_{ij}}), \quad ij > 1, \quad b^j \equiv c \equiv 0,$$

we can take $\lambda_0 = 0$ in i) and ii).

We will only prove Theorem 6.9. The proof of Theorem 6.8 is similar and actually simpler. We make a few preparations before the proof.

Denote

$$P_0 u = -u_t + a^{ij} D_{ij} u,$$

where $a^{11} = a^{11}(x^1)$ and $a^{ij} = a^{ij}(t, x^1, \dots, x^{\pi_{ij}})$ for $ij > 1$. We first establish a solvability result for equations with these simple leading coefficients.

Theorem 6.11. *Let $p \in (1, \infty)$ and $T \in (-\infty, \infty]$. Then for any $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ and $\lambda \geq 0$, we have*

$$\begin{aligned} \lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} + \|D^2u\|_{L_p(\mathbb{R}_T^{d+1})} + \|u_t\|_{L_p(\mathbb{R}_T^{d+1})} \\ \leq N \|P_0u - \lambda u\|_{L_p(\mathbb{R}_T^{d+1})}, \end{aligned} \quad (6.4)$$

where $N = N(d, p, \delta) > 0$. Moreover, for any $f \in L_p(\mathbb{R}_T^{d+1})$ and $\lambda > 0$ there is a unique $u \in W_p^{1,2}(\mathbb{R}_T^{d+1})$ solving

$$P_0u - \lambda u = f \quad \text{in } \mathbb{R}_T^{d+1}.$$

Proof. We follow closely the proof of Theorem 4.1, and recall that we only need to prove (6.4) for $u \in C_0^\infty$ and $T = \infty$. We make a change of variables as in the proof of Theorem 4.1. Denote

$$\begin{aligned} v(t, y^1, y') &= u(t, \phi^{-1}(y^1), y'), \\ \bar{a}^{11}(y^1) &= a^{11}(\phi^{-1}(y^1)), \\ \bar{a}^{ij}(t, y) &= a^{ij}(t, \phi^{-1}(y^1), y^2, \dots, y^{\pi_{ij}}), \quad ij > 1, \\ \bar{f}(t, y) &= f(t, \phi^{-1}(y^1), y'). \end{aligned}$$

Define a divergence form operator \bar{P}_0 in the (t, y) -coordinates by

$$\begin{aligned} \bar{P}_0v &= -v_t + D_1 \left(\frac{1}{\bar{a}^{11}} D_1 \bar{v} \right) + \sum_{j=2}^d D_j \left(\frac{\bar{a}^{1j} + \bar{a}^{j1}}{\bar{a}^{11}} D_1 v \right) \\ &+ \sum_{1 < i < j} D_j ((\bar{a}^{ij} + \bar{a}^{ji}) D_i v) + \sum_{i=2}^d D_i (\bar{a}^{ii} D_i v). \end{aligned}$$

Clearly, v satisfies in \mathbb{R}^{d+1}

$$\bar{P}_0v - \lambda v = \bar{f}. \quad (6.5)$$

This is a divergence form equation, which satisfies the condition of Theorem 6.3 iii). Thus, we have

$$\lambda \|v\|_{L_p} + \sqrt{\lambda} \|Dv\|_{L_p} \leq N \|\bar{f}\|_{L_p}.$$

Therefore,

$$\lambda \|u\|_{L_p} + \sqrt{\lambda} \|Du\|_{L_p} \leq N \|f\|_{L_p}. \quad (6.6)$$

In order to estimate the second derivatives, notice that $D_d v$ satisfies

$$\bar{P}_0(D_d v) - \lambda D_d v = D_d \bar{f}.$$

Again by Theorem 6.3 iii), we get

$$\|D_{yy^d} v\|_{L_p} \leq N \|\bar{f}\|_{L_p},$$

which implies

$$\|D_{xx^d} u\|_{L_p} \leq N \|f\|_{L_p}. \quad (6.7)$$

Next, we estimate $D_{xx^{d-1}}u$. Upon moving the terms involving D_d to the right-hand side of (6.5), adding D_d^2v to both sides of (6.5), and taking one derivative in y_{d-1} on both sides, we get

$$\begin{aligned} \bar{P}_1(D_{d-1}v) - \lambda D_{d-1}v &= D_{d-1}\bar{f} - \sum_{i=1}^{d-1} D_{d-1} \left((\bar{a}^{id} + \bar{a}^{di}) D_{id}v \right) \\ &\quad + D_{d-1} \left(\frac{\bar{a}^{1d} + \bar{a}^{d1}}{\bar{a}^{11}} D_{1d}v \right) + D_{d-1}((1 - \bar{a}^{dd})D_d^2v), \end{aligned}$$

where \bar{P}_1 is defined by

$$\begin{aligned} \bar{P}_1v &= -v_t + D_1 \left(\frac{1}{\bar{a}^{11}} D_1\bar{v} \right) + \sum_{j=2}^{d-1} D_j \left(\frac{\bar{a}^{1j} + \bar{a}^{j1}}{\bar{a}^{11}} D_1v \right) \\ &\quad + \sum_{1 < i < j < d} D_j ((\bar{a}^{ij} + \bar{a}^{ji}) D_i v) + \sum_{i=2}^{d-1} D_i (\bar{a}^{ii} D_i v) + D_d^2v. \end{aligned}$$

By Theorem 6.3 iii), we get

$$\|D_{yy^{d-1}}v\|_{L_p} \leq N\|\bar{f}\|_{L_p} + N\|D_{yy^d}v\|_{L_p},$$

which together with (6.7) implies

$$\|D_{xx^{d-1}}u\|_{L_p} \leq N\|f\|_{L_p}.$$

We can repeat this procedure $d-1$ times, and then obtain

$$\|D_{xx'}u\|_{L_p} \leq N\|f\|_{L_p}. \quad (6.8)$$

Finally, from (3.2), we see that $w := D_1u$ satisfies

$$-w_t + D_1(a^{11}D_1w) + \Delta_{d-1}w - \lambda w = D_1f + \sum_{ij>1} D_1((\delta_{ij} - a^{ij})D_{ij}u).$$

We use Theorem 6.3 iii) one more time to get

$$\|D_1^2u\|_{L_p} \leq \|Dw\|_{L_p} \leq N\|f\|_{L_p} + N \sum_{ij>1} \|D_{ij}u\|_{L_p}. \quad (6.9)$$

Combining (6.6), (6.8), (6.9) and (4.6) yields (6.4). The theorem is proved. \square

In the sequel, we again only consider the case $d = 3$, or $d > 3$ and π_{ij} is replaced by $\tilde{\pi}_{ij}$ in Assumption 6.2. Like before, the general case follows by an induction.

Theorem 6.12. *Let $d \geq 2$, $p \in (1, \infty)$, $\gamma > 0$, $\tau, \sigma \in (1, \infty)$ satisfying $1/\tau + 1/\sigma = 1$. Assume $b^i = c = 0$ and $u \in W_p^{1,2}$. Then under Assumption 6.2 (γ) with $\tilde{\pi}_{ij}$ in place of π_{ij} , there exist positive constants $\alpha = \alpha(d, \delta) \in (0, 1)$*

and N depending only on d , p , δ , and τ such that, for any $(t_0, x_0) \in \mathbb{R}^{d+1}$, $r \in (0, \infty)$, and $\kappa \geq 4$,

$$\begin{aligned} \sum_{i=3}^d (|D_i^2 u(t, x) - (D_i^2 u)_{Q_r(t_0, x_0)}|^p)_{Q_r(t_0, x_0)} &\leq N \kappa^{d+2} (|Pu|^p)_{Q_{\kappa r}(t_0, x_0)} \\ &+ N \kappa^{d+2} \gamma^{1/\sigma} (|D^2 u|^{p\tau})_{Q_{\kappa r}(t_0, x_0)}^{1/\tau} + N \kappa^{-p\alpha} \sum_{i=3}^d (|D_i^2 u|^p)_{Q_{\kappa r}(t_0, x_0)}, \end{aligned}$$

provided that u vanishes outside Q_{R_0} .

Proof. The proof is similar to that of Theorem 3.5 by using Theorem 6.11 instead of Theorem 3.1. We omit the details. \square

Lemma 6.13. *Assume $b^i = c = 0$. For any $p \in (1, \infty)$ there exists a γ_2 , μ_2 and N , depending only on d , δ and p , such that under Assumption 6.2 (γ_2) with $\tilde{\pi}_{ij}$ in place of π_{ij} , for any $u \in C_0^\infty(Q_{\mu_2^{-1}R_0})$ we have*

$$\|D^2 u\|_{L_p} + \|u_t\|_{L_p} \leq N \|Pu\|_{L_p} + N \sum_{i=3}^d \|D_i^2 u\|_{L_p}. \quad (6.10)$$

Proof. Set $f = Pu$. Note that u satisfies

$$-u_t + \sum_{i,j=1}^2 a^{ij} D_{ij} u + \sum_{i=3}^d D_i^2 u = f + \sum_{\max(i,j)>2} (\delta_{ij} - a^{ij}) D_{ij} u.$$

The coefficients on the left-hand side above satisfy Assumption 2.1 (γ_2) ii) and iii). Thus by Theorem 4.2 and Lemma 4.3, for γ_2 sufficiently small and μ_2 sufficiently large depending only on d , δ and p , we have

$$\|u_t\|_{L_p} + \|D^2 u\|_{L_p} \leq N \|f\|_{L_p} + N \sum_{\max(i,j)>2} \|D_{ij} u\|_{L_p}.$$

By an inequality similar to (3.9) we reach (6.10). The lemma is proved. \square

Proof of Theorem 6.8. Assume for the moment that in Assumption 6.2 π_{ij} is replaced by $\tilde{\pi}_{ij}$. We first establish a counterpart of Lemma 3.7 by using Theorem 6.12 and Lemma 6.13. After that we prove the theorem by arguing as in the proof of Theorem 2.2. The general case follows by an induction. The theorem is proved. \square

APPENDIX A. ESTIMATES FOR DIVERGENCE FORM EQUATIONS

For reader's convenience, in the Appendix section, we recall a few results obtained in [7] and [10] for divergence form parabolic equations and systems. These results have been used in this paper.

First, consider parabolic equations in divergence form:

$$-u_t + D_j(a^{ij} D_i u) - \lambda u = \operatorname{div} g + f, \quad (\text{A.1})$$

where $\lambda \geq 0$ is a constant and $g = (g^1, g^2, \dots, g^d)$. We assume the same boundedness and ellipticity condition on a^{ij} as in the introduction.

Proposition A.1 (Corollary 5.5 of [7]). *Let $p \in (1, \infty)$, $T \in (-\infty, \infty]$, and $f, g \in L_p(\mathbb{R}_T^{d+1})$. Assume that $a^{11} = a^{11}(x^1)$ and $a^{ij} = a^{ij}(t, x^1)$ for $ij > 1$. Then there exists a constant $N > 0$, depending only on d, p and δ , such that for any $u \in \mathcal{H}_p^1(\mathbb{R}_T^{d+1})$ satisfying (A.1) we have*

$$\lambda \|u\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|Du\|_{L_p(\mathbb{R}_T^{d+1})} \leq N \sqrt{\lambda} \|g\|_{L_p(\mathbb{R}_T^{d+1})} + N \|f\|_{L_p(\mathbb{R}_T^{d+1})},$$

provided that $\lambda \geq 0$. In particular, when $\lambda = 0$ and $f \equiv 0$, we have

$$\|Du\|_{L_p(\mathbb{R}_T^{d+1})} \leq N \|g\|_{L_p(\mathbb{R}_T^{d+1})}.$$

The following result follows directly from the proof of Theorem 2.5 [7].

Proposition A.2. *Let $p \in (1, \infty)$, $\lambda = 0$, $f \equiv 0$ and $g \in L_p$. Then there exist constants γ_1 , N and $\mu_1 > 0$ depending only on d, p , and δ , such that under Assumption 2.1 (γ_1) ii) and iii) the following holds. For any $u \in C_0^\infty$ vanishing outside $Q_{\mu_1^{-1}R_0}$ and satisfying (A.1) in \mathbb{R}^{d+1} , we have*

$$\|Du\|_{L_p} \leq N \|g\|_{L_p}.$$

Next, we consider parabolic systems in the form

$$- \mathbf{u}_t + D_\alpha(A^{\alpha\beta} D_\beta \mathbf{u}) - \lambda \mathbf{u} = \operatorname{div} \mathbf{g}, \quad (\text{A.2})$$

where $\mathbf{u} = (u^1, \dots, u^m)^T$,

$$\mathbf{g}_\alpha = (g_\alpha^1, \dots, g_\alpha^m)^{\text{tr}}, \quad \alpha = 1, \dots, d,$$

and $A^{\alpha\beta} = [A_{ij}^{\alpha\beta}]_{i,j=1}^d$ satisfies (5.1). Moreover, we suppose

$$A^{11} = A^{11}(t), \quad A^{\alpha\beta} = A^{\alpha\beta}(t, x^1), \quad \alpha\beta > 1.$$

The proposition below is a special case of Theorem 5.1 [10].

Proposition A.3. *Let $p \in (1, \infty)$, $\lambda \in (0, \infty)$, $T \in (-\infty, \infty]$ and $\mathbf{g} \in L_p$. Suppose that $\mathbf{u} \in \mathcal{H}_p^1(\mathbb{R}_T^{d+1})$ satisfies (A.2) in \mathbb{R}_T^{d+1} . Then,*

$$\|D\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} + \sqrt{\lambda} \|\mathbf{u}\|_{L_p(\mathbb{R}_T^{d+1})} \leq N \|\mathbf{g}\|_{L_p(\mathbb{R}_T^{d+1})},$$

where $N = N(d, m, \delta, p)$.

The following Hölder estimate is proved in [10] by using a bootstrap argument.

Proposition A.4 (Lemma 6.3 [10]). *Let $p \in (1, \infty)$. Assume $\mathbf{u} \in C_{loc}^\infty$ satisfies (A.2) in Q_2 with $\lambda = 0$ and $\mathbf{g} \equiv 0$. Then for any $\gamma \in (0, 1)$, we have*

$$\|\mathbf{u}\|_{C^{\gamma/2, \gamma}(Q_1)} + \|D_{x'} \mathbf{u}\|_{C^{\gamma/2, \gamma}(Q_1)} \leq N \|\mathbf{u}\|_{L_p(Q_2)},$$

where $N = N(d, m, \delta, \gamma, p)$.

ACKNOWLEDGEMENT

The author is grateful to Nicolai V. Krylov and Doyoon Kim for very helpful discussions and comments during the preparation of this paper. He also would like to thank the referee and the associate editor for their helpful comments on an earlier version of the paper.

REFERENCES

- [1] E. Acerbi, G. Mingione, Gradient estimates for a class of parabolic systems, *Duke Math. J.* **136** (2007), no. 2, 285–320.
- [2] M. Bramanti, M. Cerutti, $W_p^{1,2}$ solvability for the Cauchy-Dirichlet problem for parabolic equations with VMO coefficients, *Comm. Partial Differential Equations* **18** (1993), no. 9-10, 1735–1763.
- [3] S. Byun, L. Wang, Gradient estimates for elliptic systems in non-smooth domains, *Math. Ann.* **341** (2008), no. 3, 629–650.
- [4] F. Chiarenza, M. Frasca, P. Longo, Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients, *Ricerche Mat.* **40** (1991), no. 1, 149–168.
- [5] ———, $W^{2,p}$ -solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients, *Trans. Amer. Math. Soc.* **336** (1993), no. 2, 841–853.
- [6] G. Chiti, A $W^{2,2}$ bound for a class of elliptic equations in nondivergence form with rough coefficients, *Invent. Math.* **33** (1976), no. 1, 55–60.
- [7] H. Dong, Solvability of parabolic equations in divergence form with partially BMO coefficients, *J. Funct. Anal.* **258** (2010), no. 7, 2145–2172..
- [8] H. Dong, D. Kim, Parabolic and elliptic systems with VMO coefficients, *Methods Appl. Anal.* **16** (2009), no. 3, 365–388.
- [9] ———, Elliptic equations in divergence form with partially BMO coefficients, *Arch. Ration. Mech. Anal.* **196** (2010), no. 1, 25–70.
- [10] ———, Parabolic and elliptic systems in divergence form with partially BMO coefficients, *Calc. Var. Partial Differential Equations*, to appear (2010), DOI: 10.1007/s00526-010-0344-0.
- [11] H. Dong, N. V. Krylov, Second-order elliptic and parabolic equations with $B(\mathbb{R}^2, VMO)$ coefficients, *Trans. Amer. Math. Soc.* **362** (2010), no. 12, 6477–6494.
- [12] D. Gilbarg, N. S. Trudinger, “Elliptic partial differential equations of second order”, Reprint of the 1998 edition, Classics in Mathematics, Springer-Verlag, Berlin, 2001.
- [13] R. Haller-Dintelmann, H. Heck, M. Hieber, L^p - L^q -estimates for parabolic systems in non-divergence form with VMO coefficients, *J. London Math. Soc. (2)* **74** (2006), no. 3, 717–736.
- [14] D. Kim, Parabolic equations with measurable coefficients. II. (English summary) *J. Math. Anal. Appl.* **334** (2007), no. 1, 534–548.
- [15] ———, Elliptic and parabolic equations with measurable coefficients in L_p -spaces with mixed norms, *Methods Appl. Anal.* **15** (2008), no. 4, 437–467.
- [16] ———, Parabolic equations with partially BMO coefficients and boundary value problems in Sobolev spaces with mixed norms, *Potential Anal.* **33** (2010), no. 1, 17–46.
- [17] D. Kim, N. V. Krylov, Elliptic differential equations with coefficients measurable with respect to one variable and VMO with respect to the others, *SIAM J. Math. Anal.* **39** (2007), no. 2, 489–506.
- [18] ———, Parabolic equations with measurable coefficients, *Potential Anal.* **26** (2007), no. 4, 345–361.
- [19] P.C. Kunstmann, On maximal regularity of type L^p - L^q under minimal assumptions for elliptic non-divergence operators, *J. Funct. Anal.* **255** (2008), no. 10, 2732–2759.

- [20] N. V. Krylov, “Nonlinear elliptic and parabolic equations of second order”, Nauka, Moscow, 1985 in Russian; English translation Reidel, Dordrecht, 1987.
- [21] ———, Parabolic and elliptic equations with VMO coefficients, *Comm. Partial Differential Equations* **32** (2007), no. 3, 453–475.
- [22] ———, Parabolic equations with VMO coefficients in Sobolev spaces with mixed norms, *J. Funct. Anal.* **250** (2007), no. 2, 521–558.
- [23] ———, Second-order elliptic equations with variably partially VMO coefficients, *J. Funct. Anal.* **257** (2009), no. 6, 1695–1712.
- [24] G. Lieberman, “Second order parabolic differential equations”, World Scientific, Singapore-New Jersey-London-Hong Kong, 1996.
- [25] A. Lorenzi, On elliptic equations with piecewise constant coefficients, *Applicable Anal.* **2** (1972), 79–96.
- [26] A. Maugeri, D. Palagachev, L. Softova, “Elliptic and parabolic equations with discontinuous coefficients”, Math. Res., vol. 109, Wiley—VCH, Berlin, 2000.
- [27] D. Palagachev, L. Softova, Characterization of the interior regularity for parabolic systems with discontinuous coefficients, *Atti. Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* **16** (2005), 125–132.

(H. Dong) DIVISION OF APPLIED MATHEMATICS, BROWN UNIVERSITY, 182 GEORGE STREET, PROVIDENCE, RI 02912, USA

E-mail address: Hongjie.Dong@brown.edu